UNIVERSIDAD CARLOS III DE MADRID

working papers

UC3M Working Papers
Statistics and Econometrics
17-03
ISSN 2387-0303
Abril 2017

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-48

Evaluating significant effects from alternative seeding systems: A Bayesian approach, with an application to the UEFA Champions League

Francisco Corona[a], David Forrest[c], J.D. Tena[c,d] and Michael Wiper[a,b]

Abstract

The paper discusses how to evaluate alternative seeding systems in sports competitions. Prior papers have developed an approach which uses a forecasting model at the level of the individual match and then applies Monte Carlo simulation of the whole tournament to estimate the probabilities associated with various outcomes or combinations of outcomes. This allows, for example, a measure of outcome uncertainty to be attached to each proposed seeding regime. However, this established approach takes no note of the uncertainty surrounding the parameter estimates in the underlying match forecasting model and this precludes testing for statistically significant differences between probabilities or outcome uncertainty measures under alternative regimes. We propose a Bayesian approach which resolves this weakness in standard methodology and illustrate its potential by examining the effect of seeding rule changes implemented in the UEFA Champions League, a major football tournament, in 2015. The reform appears to have increased outcome uncertainty. We identify which clubs and which sorts of clubs were favourably or unfavourably affected by the reform, distinguishing effects on probabilities of progression to different phases of the competition.

Keywords: OR in sports; seeding; football; Monte Carlo simulation; Bayesian.

[a] Department of Statistics, Universidad Carlos III de Madrid.
[b] Instituto Flores de Lemus, Universidad Carlos III de Madrid.
[c] Management School, University of Liverpool.
[d] Università di Sassari and CRENoS.

Acknowledgements: the fourth author acknowledge financial support from the Spanish Ministry of Education and Science, research project ECO2012-32401.

Evaluating significant effects from alternative seeding systems: A Bayesian approach, with an application to the UEFA Champions League

Francisco Corona[1], David Forrest[2], J.D. Tena[3] and Michael Wiper[1]

[1]Department of Statistics, Universidad Carlos III de Madrid
[2]Management School, University of Liverpool
[3]University of Liverpool (UK) and Università di Sassari and CRENoS (Italy)

April 2017

Abstract

The paper discusses how to evaluate alternative seeding systems in sports competitions. Prior papers have developed an approach which uses a forecasting model at the level of the individual match and then applies Monte Carlo simulation of the whole tournament to estimate the probabilities associated with various outcomes or combinations of outcomes. This allows, for example, a measure of outcome uncertainty to be attached to each proposed seeding regime. However, this established approach takes no note of the uncertainty surrounding the parameter estimates in the underlying match forecasting model and this precludes testing for statistically significant differences between probabilities or outcome uncertainty measures under alternative regimes. We propose a Bayesian approach which resolves this weakness in standard methodology and illustrate its potential by examining the effect of seeding rule changes implemented in the UEFA Champions League, a major football tournament, in 2015. The reform appears to have increased outcome uncertainty. We identify which clubs and which sorts of clubs were favourably or unfavourably affected by the reform, distinguishing effects on probabilities of progression to different phases of the competition.

Keywords: OR in sports; seeding; football; Monte Carlo simulation; Bayesian.

1 Introduction

Across a range of sports, seeding is commonly employed to influence which players or teams compete against each other in the successive stages of an elimination tournament. Typically, the arrangements are designed to prevent the competitors judged ex ante to be the strongest from playing each other in the initial rounds and knocking each other out. In turn, this will increase the probability that the final and most prestigious stages will feature only the best players, raising interest (and, potentially, commercial revenue) on account of both the quality of players on show and the degree of competitive balance.

But seeding can go too far in terms of how rigorously it is applied. If the early rounds are contrived to feature only matches between the strongest and the weakest, they may become too 'predictable', attracting little interest from spectators and television viewers, who may indeed become disillusioned with the whole competition.

Different seeding systems are possible. Sports organisers have to take the decision on which is the most satisfactory, trading off perhaps between the gains to be had from delivering the top players or clubs to the finals and the commercial loss if the early rounds are too formulaic and lacking in suspense. Their decisions can in practice be helped by the application of 'sport analytics', a term interpreted by Wright (2014) as referring to "sporting applications of Operational Research".

The present paper has two purposes, one specific and the other more general. Our specific goal is to analyse the effects of seeding arrangements in the UEFA Champions League, a football tournament whose Final was described by Chadwick (2015) as the richest sporting event in the World. From season 2015-6, the organisers applied a new seeding system to the tournament. We will compare the old and new seeding rules (and also a hypothetical open-draw system with no seeding) in terms of the probabilities of participating clubs reaching various stages of the competition. This will help identify which sorts of clubs were gainers and which losers from the reform in terms of probabilities of progression to each successive stage of the competition. This matters for financial as well as sporting reasons. For example, in the 2015-6 edition of the Champions League, each club moving into the round of 16 was awarded €5.5m. Prize money increases with each extra round. By the time it claimed the trophy, Real Madrid had earned prize money of €80.07m and the beaten finalists Atlético de Madrid received €69.66m. Each round a club survives also provides an opportunity to gain additional revenue in ticket sales from another high profile game. In some European leagues, the sums that clubs could generate from advancing deep into the competition could be sufficient to give them the extra resources for new players that will allow them to dominate their domestic league in the following season.

In addition to focusing on how the change in rules created advantages for particular clubs in terms of raising their chances of winning substantial prize money, we consider also the impact on the attractiveness of the competition as a whole, employing an entropy measure to capture the degree of outcome uncertainty. In particular, we compare uncertainty of outcome, between alternative seeding arrangements, over the identity of the eventual winner.

The second and more general purpose of the paper is to refine methodology developed in prior papers to allow us to establish a template for assessing future proposed changes in seeding systems and other features of tournament design across sports, including in the UEFA Champions League itself. For example, from season 2018-9, there will be changes in the Champions League not in the seeding arrangements as such but in other rules such as the distribution of places in the competition between member countries. Our methods could readily be adapted to allow assessment of such reforms in terms of their effect on the predictability of the tournament.

Our methodology is related to a strand of literature (for example Koning et al., 2003, Suzuki et al., 2010) in which simulation probability models are constructed to identify which teams have the greatest probability of winning a championship. Scarf and Yusof (2011) extended this approach in the context of the 2010 FIFA World Cup by using simulation to derive probabilities of success for the competing teams not only for the tournament design which was actually in place but also for hypothetical contests where the same teams took part but with alternative seeding arrangements. Since they used the results to compare outcome uncertainty under alternative seeding regimes, their paper was a natural starting point for developing our approach.

The possibility of employing simulation to examine and estimate the change in seeding arrangements in the Champions League also attracted Dagaev and Suzdaltsev (2016). Their focus of interest were different from ours, concentrating on the competitive balance of the Final of the competition and providing extended treatment of the probabilities of different countries rather than individual clubs winning the championship. By contrast, our contribution looks at probabilities of clubs progressing to each phase of the competition.

The methodology followed by Dagaev and Suzdaltsev (2016) also differed from ours in that they, like the earlier papers cited above, estimated a probabilistic forecasting model for individual football matches based on some measure or measures of the strength of each team. Each of these papers then applied Monte Carlo simulation to the whole tournament, where probabilities of outcomes of each game were represented by point estimates derived from the individual match forecasting model.

However, these match-level probabilities themselves are subject to uncertainty because the point parameters are just one observation from a distribution. Ignoring this uncertainty will not affect the efficacy of the approach if the purpose is simply to generate a point estimate of each team's chance of winning the competition. But our purpose is to compare probabilities of success under alternative seeding regimes. For example, it might be that, following the standard approach adopted by previous authors, it was determined that, for a particular club, the probability of its winning the tournament would be 0.10 under one seeding regime and 0.15 under the other. But are these probability estimates statistically significantly different from each other? With the standard approach used in previous papers, it would be impossible to say. However, with our Bayesian approach, it is possible to obtain the posterior distribution of the parameters in the model and then run a different Monte Carlo simulation for each of the points of the distribution, weighted by their relative frequencies. This would allow us to get a more complete picture of the winning probabilities. Using this information, we can decide whether differences in the probabilities of winning the tournament are statistically significant.

Similarly, we estimate the uncertainty of outcome regarding which club will finally win the trophy. Rather than just present two point estimates of a measure of outcome uncertainty, our methodology will permit us to say whether alternative seeding regimes offer statistically significantly different degrees of outcome uncertainty.

The paper proceeds as follows. In section 2, we describe the structure of the UEFA Champions League and the changes made to the seeding rules in 2015-6. Section 3 describes the entropy measure we will employ to assess the degree of outcome uncertainty over which clubs will reach various stages of the competition (including winning it). In section 4, we set out our Bayesian Poisson model for probabilistic forecasting of individual match results and in section 5 we describe the execution and results of the Monte Carlo simulation. Section 7 consists of conclusions and reflections.

2 Seeding rules in the UEFA Champions League

The structure of the competition was not changed when the new seeding rules were introduced. It remained the case that 32 clubs participated in each annual edition of the tournament. One of these places was reserved for the current holders of the Champions League trophy. Twenty-one places were for clubs from relatively strong footballing countries which qualified automatically through their finishing positions in their domestic leagues. The final ten places were for clubs which were successful in a qualifying tournament held in the summer (where participation in the qualifiers was again conditional on finishing positions in domestic leagues).

Each year, by the end of August, the names of the 32 clubs taking part are known. At this stage, the competitive balance of the tournament (the variation in strength across the competing clubs) is fixed. But the outcome uncertainty of the tournament (the variation between clubs in the probabilities of winning) will be influenced by the seeding arrangements applied.

According to the taxonomy of tournament designs set out in Scarf and Yusof (2011), the Champions League is a "hybrid 1G-KO" competition. This means that, just as in the World Cup, there is a single "group stage" (played September to December) followed by a series of knock-out or elimination rounds (February to May) culminating in a grand final (June).

At the start of the competition, the 32 clubs are split into eight groups (A to H). Each group is a mini double round-robin league such that each club plays each other club both home and away. A win is rewarded with three points and a draw with one point. The top two clubs in each group proceed to the round of 16. From then on, the tournament is a straight knock-out competition with a fresh draw conducted for each round. However, a constraint imposed on the draw for the round of 16 is that top place finishers from the group stage play second-placed finishers from the group stage.

The point at which seeding is applied is in the draw which allocates clubs between the eight initial groups. For the draw, clubs are allocated to either pot 1, pot 2, pot 3 or pot 4 according to their seeding. The idea is that each group will then include one team from each pot. A slight complication to be noted is that each group is also constrained to consist of clubs from four different countries.

Seeding is based primarily on the "UEFA coefficients" of each club prior to the tournament.

These are ranking points earned by wins and progression in the UEFA Champions League and in the UEFA Europa League in the preceding five seasons by the club itself and (with lesser weight) by clubs from the national association to which it belongs.

Prior to season 2015-6, allocation of clubs was according to UEFA coefficients. Thus pot 1 consisted of the eight strongest clubs according to UEFA rankings and none of these could play each other in the group stage. This is classic seeding which is intended to minimise the probability of an early exit by one of the "best" competitors. Similarly clubs were distributed across the other pots in order according to their ranking by UEFA coefficient.

From season 2015-6, different seeding arrangements were put into place. Pot 1 would now be populated by the reigning champion of the Champions League and the champion clubs from the seven strongest national leagues according to UEFA national coefficients. For that particular edition of the Champions League, Barcelona was entitled to a place in pot 1 both from being Champions League and Spanish League champion in the previous season, so an extra place was available for the champion club from the eighth ranked national league (Netherlands). All the remaining clubs were then distributed across pots 2, 3 and 4 according to rank order of coefficient rankings.

This apparently minor change in the seeding rules for allocation to pot 1 had the potential to affect the degree of outcome uncertainty surrounding the tournament since it did shift clubs between groups. For example, according to the UEFA coefficients, Real Madrid was the strongest club of all but it was in pot 2 because it had failed to win the Spanish League. It is hard from intuition to be confident about how this will change outcome uncertainty. For example, it might reduce the probability of progression for clubs like Chelsea (which would have been in pot 1 under either system) since Chelsea is not now protected from the possibility of playing the mighty Real Madrid in the group stage. On the other hand, Chelsea might be confident of finishing in the top two even in a group with Real Madrid and then it would be protected from playing Real Madrid in the round of 16. Clearly the effects of the seeding change are hard to work out and analysis therefore requires simulation of the tournament.

3 Entropy measure

We will measure outcome uncertainty from the perspective of the point in time just prior to the draw to allocate the 32 competing clubs between the eight groups which comprise the first stage of the tournament. It is at this stage that the seeding is implemented which influences the flow of events throughout the rest of the competition.

Entropy is a measure of unpredictability of information content. It has been employed to capture uncertainty in sports since Horowitz (1997) used it to investigate the apparent "improvement" in uncertainty over time in Major League Baseball. Here, in the context of the Champions League, we are interested in measuring uncertainty over several outcomes, for example which club will win the tournament and which clubs will emerge from the group stage to take places in the Round of 16.

To illustrate for the case of uncertainty over which club will win the tournament, let $p_{j0} =$

$P(V_j|\xi_0)$ be the probability of victory for club j conditional on pre-tournament information concerning the 32 clubs in the competition.

Entropy is defined as follows:

$$e_0 = -\sum_{j=1}^{32} p_{j0}\log_2 p_{j0}. \tag{1}$$

Minimum entropy or maximum information occurs when some $p_{j0} = 1$ while the others are 0. In this case, $e_0 = 0$. At the other extreme, maximum entropy is when all probabilities p_{j0} are equal to 1/32 so that there is maximum outcome uncertainty.

Note again that entropy is calculated before the start of the tournament. Geenens (2014) and Corona et al. (2017) estimate the entropy in different phases of sports tournaments. For each, their focus is on identifying the decisive matches in a tournament. But our interest in the entropy measure is not to identify decisive matches in a given tournament. Rather it is to compare tournament outcome uncertainty under alternative seeding policies.

4 Bayesian Poisson Model

The first requirement for simulating a tournament is a viable model which can be used for probabilistic forecasting of individual matches which might take place within the tournament. Two categories of model have been popular in the academic literature. Economists (such as Koning, 2000) have tended to favour ordered probit models which directly generate probability estimates for match outcome (home win, draw, away win). Statisticians (since Dixon and Coles, 1997) have typically focused on variants of Poisson regression which yield for each team, the estimated probabilities of its scoring 0 goals, 1 goal, 2 goals etc. Combining these probabilities enables point estimation of winning probabilities for each side.

Goddard (2005) found that the performance of each class of model was similar. Here, we adopt a Poisson regression approach because, sometimes, progression in the Champions League depends on goals, not just on match outcomes. Thus, if clubs are tied in points in a group at the end of the group stage, the tie break rules include comparison of goal difference. The model underlying the simulation must therefore generate probabilities of each possible scoreline in a match as well as win-draw-lose probabilities.

To this end, we adopt a Poisson regression model (BP) for the goals scored by each team in a match as follows:

$$Y_{t,k} \sim \text{Poisson}(\lambda_{T,k}),$$

where $Y_{t,k}$ represents the number of goals scored by team T in a match at time k and

$$\log(\lambda_{T,k}) = \beta_{A_T}x_{A_T,k} + \beta_{A_O}x_{A_O,k} + \beta_{H_T}x_{H_T,k} + \beta_{Aw_T}x_{Aw_T,k} + \beta_{F_T}x_{F_T,k}, \tag{2}$$

where $x_{A_T,k}$ represents the strength of team T and $x_{A_O,k}$ is the strength of the opposing club. $x_{H_T,k}$ indicates if team T plays at home, $x_{Aw_{T,k}}$ if it plays away and and $x_{F_T,k}$ if the match is the Final of the whole competition (the only match played at a neutral ground). The parameters β_{A_T}, β_{A_O}, β_{H_T}, β_{Aw_T} and β_{F_T} are coefficients that express the relationship between the explanatory variables and $\lambda_{T,k}$.

Here, we proxy the strength of a club by its tally of UEFA points (i.e its UEFA coefficient) prior to the start of each year's competition, as recorded on the UEFA website . Such points are used to rank clubs across Europe and are earned on the basis of previous performances in UEFA competitions by the subject club and by other clubs in the national league to which it belongs. Our use of official rankings mirrors the approach of Dyte and Clarke (2000) who used a Poisson model with FIFA ranking points in their simulation of the 1998 World Cup. In some applications of the Poisson model, in particular modelling of national leagues, goals scored and conceded in past matches are used as measures of strength. This produces satisfactory forecasts (Dixon and Coles, 1997).

It may be argued that recent goals scored and conceded reflect form so that the team strength coefficients should vary over time. In the present case, the Champions League is played over a time span of eight months and simulations are from the perspective of before the start of the competition. A team's form in August, when the regular league season is just starting or about to start may not be relevant to predicting the results of possible Champions League matches in April. Thus a longer-run measure of strength, such as is provided by UEFA coefficients , seems most appropriate. Further, because of heterogeneous quality across national leagues, it is difficult to gauge relative strengths of clubs on the basis of comparing their propensities to score and concede goals in domestic competition. This again points to use of information on past European performances, which is captured by UEFA coefficients.

It would be possible to fit the Poisson regression model using a classical maximum likelihood as in Dyte and Clarke (2000) but here we prefer to use Bayesian inference. In our context, we are interested in predicting the results of future matches and an advantage of Bayesian methods as opposed to classical, plug-in based prediction is that they directly take parameter uncertainty into account.

In order to complete the Bayesian formulation, we need to define a prior distribution for the regression coefficients $\beta = (\beta_0, \beta_{A_T}, \beta_{O_T}, \beta_{H_T}, \beta_{Aw_T}, \beta_{F_T})'$ in (2). One option in the presence of good prior information would be to consider a multivariate normal prior structure for β, but here we prefer to use an improper, uniform prior distribution, $f(\beta) \propto 1$ as recommended in Martin et al. (2011).

Exact calculation of the posterior distribution under this model is impossible, but instead we can generate an (approximate) Monte Carlo sample of values from the posterior distribution using Markov chain Monte Carlo (MCMC) methods, see e.g. Robert and Casella (2004) for a good review. In particular, we use a random walk Metropolis algorithm to generate the successive elements of β from their conditional posterior distributions as in e.g. Sherlock et al. (2010).

After calculating the Monte Carlo samples, to obtain the win , draw and loss probabilities, $p_{W,k}$, $p_{D,k}$ and $p_{L,k}$ respectively, we follow the procedure used by Dyte and Clarke (2000), Suzuki et al. (2010), Corona et al. (2017), among many others. In this way, the probabilities for each game during the competition for teams T against the teams O become as follows:

$$p_{W,k} = \sum_{i_T=1}^{\infty} \sum_{i_O=1}^{i_T-1} P(y_{T,k} = i_T) P(y_{O,k} = i_O), \tag{3a}$$

$$p_{D,k} = \sum_{i_T=1}^{\infty} P(y_{T,k} = i_T) P(y_{O,k} = i_T), \tag{3b}$$

$$p_{L,k} = \sum_{i_T=1}^{\infty} \sum_{i_O=1}^{i_O-1} P(y_{T,k} = i_T) P(y_{O,k} = i_O), \tag{3c}$$

where i_T and i_O are the number of scored goals for teams T and O respectively, where $P(y_{T,k} = i_T)$ and $P(y_{O,k} = i_O)$ represent the corresponding Possion probabilities of goals scored (with $\lambda_{T,k}$ and $\lambda_{O,k}$ as means) by the teams in the game. Similar to previous authors including Scarf and Yusof (2011) in their tournament simulation, we assume independence between the goals scored by the two teams.

To estimate the model, we employed a data set containing the results of every Champions League match played between seasons 2002-3 and 2014-5. This is the period since the current 1G-KO structure of the tournament was put into place.

We sampled the posterior distribution of the parameters, β, in (2) using 10,000 MCMC iterations with 5,000 iterations to burn in the chain and thinning to reduce autocorrelation. Figure 1 shows the traces of 1000 thinned values and the associated densities of each posterior parameter distribution. We can see that all traces have a white noise behaviour which suggests that convergence has been achieved.

Analysing the quantiles of the parameters, the strength of club T, the strength of the opposing team, the home effect and the "final" effect are significant.

Table 1 summarises the central values of the sample posterior distributions. To illustrate the interpretation of these values, consider a hypothetical match in which Real Madrid played at home versus Barcelona in the 2015-6 Champions League. At the start of the competition, the UEFA coefficients of the two clubs were 172 and 165 respectively but of course Real Madrid would have enjoyed home advantage in this imaginary fixture.

Using median values of the posterior distribution of the estimated parameters, the expected number of goals for each team is then:

$$\lambda_{\text{Real Madrid}} = exp\,(0.0057(172) - 0.0050(165) + 0.3520) = 1.66,$$

$$\lambda_{\text{Barcelona}} = exp\,(0.0050(165) - 0.0057(172) + 0.0014) = 1.09.$$

Rounding to the nearest integer, the single most likely result was 2-1.

When tested, our match-level model exhibited satisfactory performance in terms of predictive power. We followed a cross-validation approach, taking.the (chronologically) first 70% of matches in the data set as the training set. Thus we estimated the parameters of the model on 962 matches in the Champions League.

Having estimated the parameters of the model from the training sample, we then predicted match results in the test sample (the remaining 413 matches) using the median of the sample posterior distribution. In terms of the win/ draw/ loss outcome of each match, the most likely outcome from the model coincided with the actual outcome in 50.7% of matches.

In addition, we used the model estimated over the sample period 2003-15 to predict the home/ draw/ loss outcome of all matches which took place in the Champions League, 2015-16. In this out-of-sample testing, we again took the most likely outcome indicated by the model as our prediction and this time the success-rate was 51.2%. This is close to the success-rate (51.6%) reported by Geenens (2014) for kernel regression modelling applied to one edition of the European Championships (a 1G-K0 football tournament for national teams). Geenens (2014) notes that his was a superior success rate to that reported in some earlier literature (albeit is hard to compare performance across competitions which differ in both format and heterogeneity of team strength). We were therefore content to use our match-level model in the simulation.

5 Monte Carlo Simulation

We carried out three separate simulations of the 2015-6 Champions League. Two of these related to the old and new seeding regimes described in section 2 above. The third simulation assumed a completely random draw to determine the allocation of the 32 clubs between the 8 initial groups. We term these three possible seeding systems as the traditional (denoted by T), the new (denoted by N) and the random (denoted as R).

Koning et al. (2003) notes that it is important that procedures in a simulation exercise mimic the tournament in terms of reflecting the competition rules. In Koning's case, this was a little more straightforward than in our case. He was studying the FIFA World Cup and forecasting from a point subsequent to the allocation of national teams between the groups in the initial group stage. By contrast, our analysis is from the perspective of the time prior to the draw for the group stage. Further, there is no fresh draw in the FIFA World Cup after the group stage. All pairings in the first knock-out round are pre-determined, e.g. winners of Group A versus runners-up of Group B. By contrast there is a fresh draw for the Round of 16 in the Champions League which randomly assigns clubs from the set of group winners with clubs from the set of runners-up (with the additional constraint that clubs from the same country cannot play each other at this stage). Thus, in our case, simulation includes simulation of both the group draw and the draw for the Round of 16 (and indeed the draws for the quarter- and semi-finals) as well as simulation of the evolution of winners and losers as the tournament progresses. Similarly, to mimic the tournament closely, we apply the same tie-break rules as those set down by UEFA for determining which club proceeds to the next stage if two clubs are tied on points in a group. In the subsequent knock-out rounds, ties up to and including the semi-final are two-legged. In this case, where two teams are drawing after two legs, 30 minutes extra time is played and if the teams are still tied after extra time, a penalty shoot-out is used to decide the result. In this case, we determine the winner of the tie by a random process because of the small number of observations of extra time plus penalties in the data set.

In the context of simulation of tournaments, an innovation in how we proceed is the use of MCMC estimation (described in the previous section) instead of classical estimation. This means that we do not just obtain an estimation of the model parameters and their standard deviations but an estimation of their posterior distribution. This gives us a much fuller picture of the different values each parameter can take. Knowledge of this uncertainty can then be exploited to test the significance of differences in results from the three simulation exercises. Thus, in each of the three cases, we perform a Monte Carlo simulation for each of the 1,000 estimated parameter values in the posterior distribution.

In this section we present the results of the Monte Carlo experiments. For each club in each MCMC and considering 1,000 replications, we estimate the kernel density of the win probability using a Gaussian kernel, selecting the bandwidth according to the standard deviation of the smoothing kernel. That is, we take advantage of the fact that we estimate the whole posterior distribution of the parameters in order to appraise the significance of our results. Regarding the number of replications, we judged 1,000 adequate because, as illustrated by Figure 2, the log propensity score stabilised after around 400 or fewer replications using the mean values of the estimated parameters.

In Figure 3 we plot the density of probabilities for each p_{j0} under the traditional seeding system (corresponding average point estimates of probabilities are given in Table 2). The results are sorted in descending order for eight groups of clubs teams. It can be seen in the first panel that Real Madrid, Barcelona, Bayern Munich and Chelsea were the four clubs with the greatest chance of winning the tournament. The second group of teams with the highest probabilities of winning were Atlético de Madrid, Benfica, Porto and Arsenal. Note that, beyond these eight clubs, each of the remaining 24 had a much lower and indeed remote chance of claiming the Championship.

Figure 4 (and Table 3) plot corresponding results for the new seeding system. From the first two panels, the prospects of the top eight contenders are barely changed by the new rules. However, densities appear to be slightly more positioned to the right in the third panel, indicating that clubs like Manchester United, Valencia, Paris St. Germain and Juventus may have had their prospects a little boosted by the new seeding regime. And, from Figure 5, it appears that ending seeding altogether (system R) would have had similar effects.

With a more formal approach, we can capture the overall effect on uncertainty of outcome (in terms of which club will win the competition). Figure 6 shows the densities of the entropies for each of our three seeding regimes. The density associated with regime R is more inclined to the right than for T or N, implying that failing to seed at all will add uncertainty as to which competitor will win the tournament.

However, the magnitude of e_0 does not have any sensible meaning and it is necessary to test formally whether there are statistical differences between the entropies under each draw seeding system. Therefore we compute the confidence intervals at 95% on the differences between the means of the levels of the entropies using Tukey's Honest Significant Difference (HSD) method (Yandell, 2000). The statistic takes the following form:

$$q = \frac{\bar{x}_j - \bar{x}_i}{\sqrt{\frac{s^2}{2} 0.001}},$$

where $\bar{x}_j - \bar{x}_i$ is the difference between any pair of means and s^2 is the sample mean squared error obtained from the Analysis of Variance (ANOVA). This statistic is compared with the family wise error rate $\alpha = 0.05$ which is given by $q_{\alpha,997,3}$.

Figure 7 shows the results from three key tests for statistical significance. Together they show that the differences between seeding regimes in the level of outcome uncertainty (captured by entropy) over which club will win the tournament are statistically significant. The change from the traditional (T) to the new (N) system has increased outcome uncertainty. A random draw system with no seeding would offer greater outcome uncertainty than either.

But this is to speak only of the question of which club will win the tournament. For much of the season, interest focuses not on that question about the relatively distant future but rather on the more immediate matter of progression in the competition. In fact, the tournament features 125 individual matches each season and 96 of these are played at the group stage where what is decided is which clubs will proceed to the Round of 16. From the perspective of the sporting and commercial health of the competition, it appears important to consider whether there is any genuine suspense generated by the group stage in terms of which clubs will advance to the knock-out stage. Indeed, commentators have expressed concern that a partial explanation of falls in Champions League television viewing may be the excessive predictability of which clubs progress to the knock-out stage: "Everyone knows the competition gets going only after the turn of the year, and even moderately unexpected results in the early stages are normally evened out before the end because there are so many second chances" (The Guardian, September 15, 2016).

In addition, we should consider also how the choice of seeding regime affects the chances of individual clubs advancing to the quarter-finals and the semi-finals. Given the prize money for progression to each stage, this is important for the clubs concerned as well as for tournament organisers.

To determine the differences between the seeding systems we compute the HSD at 95% for the differences of means related to the percentage of times that each club to reaches the round of 16, the quarter-finals and the semi-finals. At this point, our Bayesian approach shows its value as it is possible to test the significance of differences between the seeding regimes.

Figure 8 shows the results for reaching the round of 16. Shakthar Donetsk, Porto, Arsenal, Benfica and Atlético de Madrid prove to have been the clubs most adversely affected by the change to the new seeding system. The common feature of these five clubs is that, historically, they performed strongly in UEFA competitions and so would previously have qualified for inclusion in pot 1. However, none won its domestic league in the preceding season and so lost top seeding and were allocated instead to pot 2.

The next group of clubs statistically significantly worse off under the new seeding rules comprises perhaps the "biggest" clubs in the competition: Chelsea, Bayern Munich, Real Madrid and Barcelona. Three of these were still in pot 1 under the new rules but all now had a reduced (though still high) probability of reaching the Round of 16. This reflects that the new system places some highly ranked clubs in pot 2 and, if two strong clubs are in the same group, there is a chance of missing progression from losing to that opposition and then making a mistake in one of the matches against the two lesser clubs.

Comfortably the club most favoured by the change in seeding regime was PSV Eindhoven. The club qualified for pot 1 as champion of the UEFA country with the eighth highest ranking. There are in fact places in pot 1 only for the champions of the seven highest rank countries (plus the Champions league winners from the previous season). However, Barcelona had won both the Champions League and the Spanish League and thus qualified on two criteria. This allowed the Dutch champion to take the resulting vacancy in pot 1 whereas, under the old system, it would have been only in pot 3. This accounts for why PSV is so strikingly an outlier in Figure 8. Similarly the next most favoured club was Juventus which, as Italian champion, qualified for pot 1 whereas previously, on the basis of its UEFA coefficient, it would have been in pot 2. Similarly, Paris St.-Germain and Zenit St. Petersburg were (statistically significant) beneficiaries of the changes according to Figure 8 and they too were able to claim places in pot 1 as national champions whereas, under the old system, each would have been in pot 2.

All these beneficiaries were more highly seeded (i.e. in a higher pot) as a result of the changes made by UEFA. But it is less straightforward to understand the gains for Sevilla and the two Manchester clubs which were in fact placed in the same pots as they would have been under the old regime (pot 3 for Sevilla, pot 2 for each of the Manchester clubs). One obvious factor is that these clubs now had a chance of being drawn into a group with a club which was highly seeded but not particularly strong according to its UEFA coefficient. This will have raised their probabilities of finishing in the top two places in their group. However, it would have been be difficult to make predictions about overall effects on these clubs' probabilities because of complications such as the effect of the rule that no group can include more than one club from any one country. Once again simulation was necessary to evaluate the broad impact of the seeding changes on overall outcome uncertainty and on the prospects of individual clubs.

Regarding progression to the quarter- and semi-finals, impacts on probabilities for each individual club are shown in Figures 9 and 10 respectively. Note that these probabilities are as assessed from the point in time immediately before the draw for the group stage is made. Hence the probabilities are not independent of the chances of survival to the Round of 16 since the latter is a precondition for the former. It is therefore unsurprising that the identities of the gainers and losers are similar to those evident from the discussion of Figure 8. At the quarter-final stage, even the ordering of gainers and losers is almost the same as at the Round of 16 qualification stage. And there are relatively minor changes in the effects of the change in seeding arrangements on the probabilities of clubs reaching the semi-finals. Together, these results indicate that any changes in clubs' prospects brought about by the regime change are being engineered mainly by impacts on probabilities of progression beyond the group stage.

The effects of the change in seeding have been illustrated here in the context of one particular edition of the Champions League, that of 2015-6. From season to season, the 32 clubs taking part will of course vary. But it is reasonable to suppose that our analysis of which particular clubs gained or lost from the reforms in 2015-6 illustrates also which types of clubs would be better or worse off in the general case. Although, for reasons of space and to avoid repetition, we do not report the details, we in fact repeated the exercise set out in this paper for the Champions League, 2014-5. That year the old seeding system was still in place but we asked what differences there would have been had the new system been in effect instead. Broadly the effects were similar to those for 2015-6 not only in terms of the difference made to entropy but also in the sorts of clubs which tended to be winners or losers. Thus, in 2014-5, CKSA Moscow would have had a significantly higher chance of progression had the new system applied: it was a national champion and would have been in pot 1 instead of pot 3 (similar to PSV Eindhoven which was the biggest gainer from regime change in 2015-6). Again, Arsenal would have had a markedly lower chance of progression had the new system been in place in 2014-5. Its previous record in UEFA competitions entitled it to a place in pot 1 whereas under the later regime it would have been in pot 2 because it had not won its national championship. In fact, Arsenal proved to be in a similar situation in 2015-6 and, as noted above, was one of the clubs most adversely affected when the new system was actually introduced.

6 Concluding remarks

The motives for seeding sports competitions may vary from case to case. Some organisers (including, we suspect, UEFA) may seek to maximise commercial revenue. Others may desire to provide the highest possible degree of outcome uncertainty or emphasise the somewhat intangible concept of "fairness" or even offer enhanced protection to the politically most powerful competitors. But, regardless of the goals of tournament organisers, the contribution of sports analytics is to enable them to compare choices over seeding arrangements in terms of the probabilities of various outcomes or combinations of outcomes occurring. Our Bayesian approach would enable richer information to be available to sports organisers evaluating changes either ex ante or ex post in seeding arrangements. In principle, the methods described could be applied across a variety of sports and tournaments. Our case study examined changes in seeding arrangements introduced into football's European Champions League in 2015. Those particular changes were found to have increased uncertainty over which club would win the competition. They also slightly reduced the probability of each of the most powerful clubs surviving to the knock-out stage while increasing the chances of some of the weakest clubs. Probably these impacts were consistent with what UEFA hoped to achieve given adverse comments on how excessively predictable the early stages of the competition had become.

References

Chadwick, S. (2015). Hard evidence: How much is the Champions league worth?. *The Conversation*, June 2, 2015.

Corona, F., Tena, J., and Wiper, M. (2017). On the Importance of the Probabilistic Model in Identifying the Most Decisive Games in a Tournament. *Journal of Quantitative Analysis in Sports*, forthcoming.

Dagaev, D. and Suzdaltsev, A. (2016). Seeding, Competitive Intensity and Quality in Knock-out Tournaments. *National Research University Higher School of Economics (HSE), Working paper.*

Dixon, M.J. and Coles, S.G. (1997). Modelling Association Footbal Scores and Inefficiencies in the Footabll Betting Market. *Journal of the Royal Statistical Society C*, 46(2):265–280.

Dyte, D. and Clarke, S.R. (2000). A ratings based Poisson model for World Cup soccer simulation. *Journal of the Operational Research Society*, 51:993–998.

Geenens, G. (2014). On the decisiveness of a game in a tournament. *European Journal of Operational Research*, 232:156–168.

Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2):331–340.

Horowitz, I. (1997). The increasing competitive balance in Major League Baseball. *Review of Industrial Organization*, 12:373–387.

Koning, R. (2000). Balance in competition in dutch soccer'. *The Statistician*, 49(3):419–431.

Koning, R., Koolhaas, M., Renes, G., and Ridder, G. (2003). A simulation model for football championships. *European Jouranl of Operational Research*, 142(2):268–276.

Martin, A., Quinn, K., and Park, J. (2011). MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software*, 43(9):1–29.

Robert, C. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer, 2 edition.

Scarf, P.A. and Yusof, M.M. (2011). A numerical study of tournament structure and seeding policy fo the soccer World Cup Finals. *Statistica Neerlandica*, 65(1):43–57.

Sherlock, C., Fearnhead, P., and Roberts, G. (2010). The Random Walk Metropolis: Linking Theory and Practice Through a Case Study. *Statistical Science*, 25(2).

Suzuki, A.K., Salasar, L.E.B., Leite, J.G., and Lozada-Neto, F. (2010). A Bayesian approach for predicting match outcomes: The 2006 (Association) Football World Cup. *Journal of the Operational Research Society*, 61:1530–1539.

Wright, M. (2014). OR analysis of sporting rules: A survey. *European Journal of Operational Research*, 232:1–8.

Yandell, B.S. (2000). *Practical Data Analysis for Designed Experiments.* Chapman & Hall.

Tables and Figures

Table 1: Summary of BP model

Variable	Median	Mean	SD
AT	0.0057	0.0057	0.0004
AO	-0.0050	-0.0050	0.0004
Home	0.3520	0.3499	0.0568
Away	0.0014	0.0002	0.0566
Final	0.3804	0.3739	0.1861

Table 2: Mean of percentages to qualify to knockout round from MCMC replications. Traditional system draw.

Team	UEFA coeff.	Round of 16	Quarter Finals	Semi Finals	Final	Winner
Arsenal	110.08	82.63	44.61	19.76	7.63	2.80
Astana	3.83	4.75	0.48	0.05	0.00	0.00
Atlético Madrid	121.00	87.32	54.28	26.89	11.74	4.75
Barcelona	165.00	97.60	78.91	55.92	37.21	22.09
BATE Borisov	35.15	13.47	2.41	0.39	0.05	0.01
Bayer Leverkusen	87.88	64.49	24.82	8.64	2.59	0.78
Bayer Munich	154.88	96.22	71.79	47.47	28.95	15.79
Benfica	118.28	86.16	49.74	24.01	10.16	4.06
Borussia Monchengladbach	33.88	13.01	2.32	0.38	0.05	0.01
Chelsea	142.08	93.86	65.59	39.24	21.32	10.44
CSKA Moscow	55.60	28.10	6.81	1.57	0.30	0.06
Dinamo Zagreb	24.70	9.80	1.46	0.21	0.03	0.00
Dynamo Kyiv	65.03	33.73	9.32	2.47	0.54	0.13
Galatasaray	50.02	25.09	5.61	1.20	0.20	0.04
Gent	13.44	6.71	0.84	0.10	0.01	0.00
Juventus	95.10	67.88	27.77	10.53	3.44	1.11
Lyon	72.98	39.96	12.41	3.65	0.89	0.23
Maccabi Tel Aviv	18.20	7.91	1.03	0.14	0.01	0.00
Malmo FF	12.54	6.51	0.79	0.09	0.01	0.00
Manchester City	87.08	63.29	23.48	8.06	2.38	0.72
Manchester United	103.08	73.37	32.92	13.67	4.92	1.73
Olympiacos	62.38	32.53	8.76	2.24	0.48	0.11
Paris Saint-Germain	100.48	72.90	31.80	12.89	4.50	1.53
Porto	111.28	83.06	44.88	20.10	7.89	2.95
PSV Eindhoven	58.20	29.93	7.57	1.81	0.35	0.07
Real Madrid	172.00	98.25	81.70	60.52	42.66	26.95
Roma	43.60	21.69	4.33	0.83	0.13	0.02
Sevilla	80.50	45.01	17.55	5.60	1.53	0.43
Shakhtar Donetsk	86.03	63.26	23.04	7.89	2.30	0.68
Valencia	100.00	73.85	35.78	14.29	4.94	1.65
Wolfsburg	31.88	12.28	2.11	0.34	0.04	0.01
Zenit Saint Petersburg	90.10	65.38	25.11	9.06	2.75	0.84

Table 3: Mean of percentages to qualify to knockout round from MCMC replications. New system draw.

Team	UEFA coeff.	Round of 16	Quarter Finals	Semi Finals	Final	Winner
Arsenal	110.08	78.90	40.49	18.28	7.20	2.68
Astana	3.83	4.87	0.53	0.06	0.00	0.00
Atlético Madrid	121.00	85.22	51.72	26.03	11.50	4.72
Barcelona	165.00	97.03	77.64	55.55	37.27	22.21
BATE Borisov	35.15	13.55	2.56	0.45	0.07	0.01
Bayer Leverkusen	87.88	65.62	26.63	9.43	2.81	0.86
Bayer Munich	154.88	94.90	69.02	46.05	28.36	15.52
Benfica	118.28	83.68	46.83	22.95	9.89	3.97
Borussia Monchengladbach	33.88	13.08	2.50	0.44	0.06	0.01
Chelsea	142.08	91.81	61.53	37.29	20.46	10.04
CSKA Moscow	55.60	28.56	7.25	1.72	0.33	0.07
Dinamo Zagreb	24.70	9.89	1.58	0.24	0.03	0.00
Dynamo Kyiv	65.03	34.22	9.99	2.69	0.61	0.14
Galatasaray	50.02	25.35	6.02	1.32	0.23	0.05
Gent	13.44	6.83	0.89	0.10	0.01	0.00
Juventus	95.10	71.01	30.93	11.86	3.92	1.27
Lyon	72.98	39.54	12.84	3.81	0.95	0.25
Maccabi Tel Aviv	18.20	8.04	1.17	0.15	0.02	0.00
Malmo FF	12.54	6.61	0.85	0.10	0.01	0.00
Manchester City	87.08	65.01	26.15	9.15	2.72	0.82
Manchester United	103.08	75.00	36.00	15.07	5.49	1.95
Olympiacos	62.38	32.63	9.28	2.39	0.51	0.12
Paris Saint-Germain	100.48	74.28	34.45	14.11	4.98	1.70
Porto	111.28	79.02	40.67	18.66	7.46	2.84
PSV Eindhoven	58.20	45.31	12.24	3.00	0.59	0.13
Real Madrid	172.00	97.62	79.94	59.68	42.33	26.82
Roma	43.60	21.91	4.68	0.93	0.15	0.03
Sevilla	80.50	47.51	19.57	6.31	1.72	0.48
Shakhtar Donetsk	86.03	48.28	18.35	6.47	1.94	0.57
Valencia	100.00	74.75	37.76	15.24	5.25	1.77
Wolfsburg	31.88	12.27	2.24	0.38	0.05	0.01
Zenit Saint Petersburg	90.10	67.69	27.74	10.09	3.12	0.95

Figure 1: Trace (left panel) and density (right) of each posterior parameter distribution of a Poisson regression model in 1,000 MCMC.

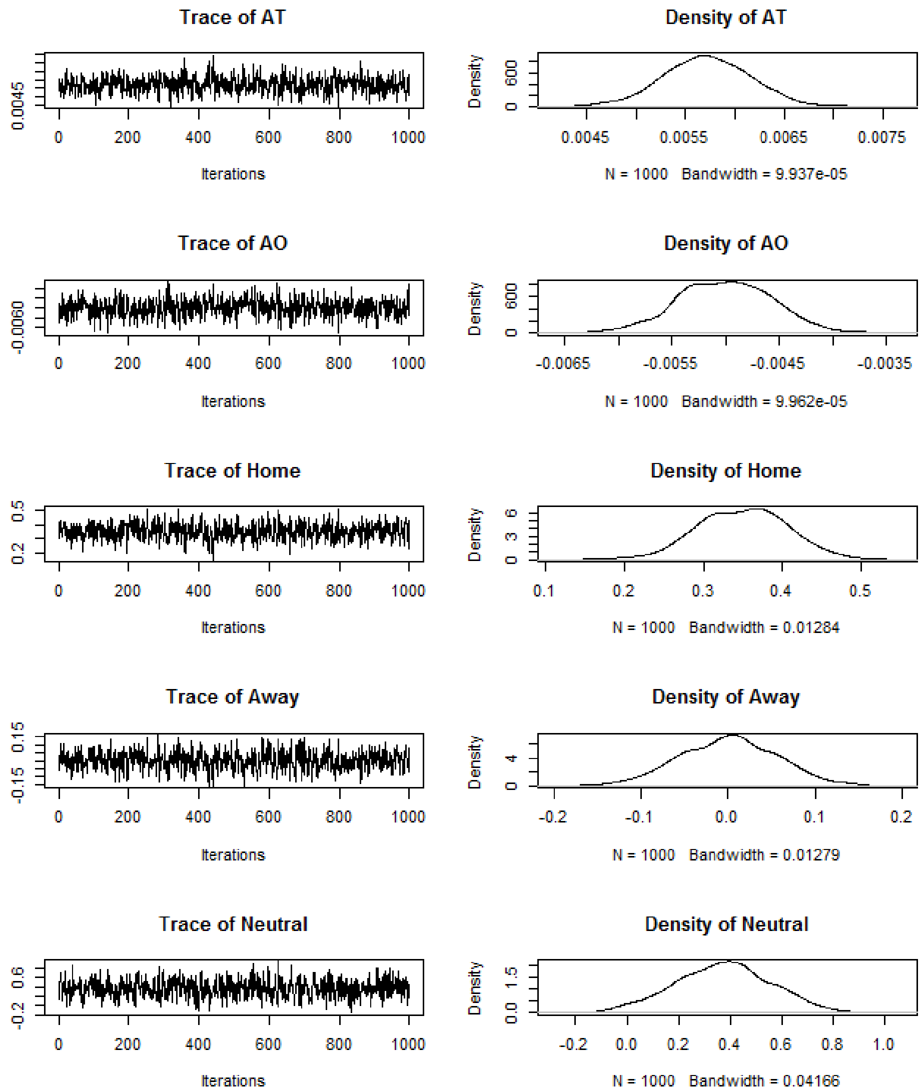

Figure 2: Running mean plots of the entropy from the MCMC

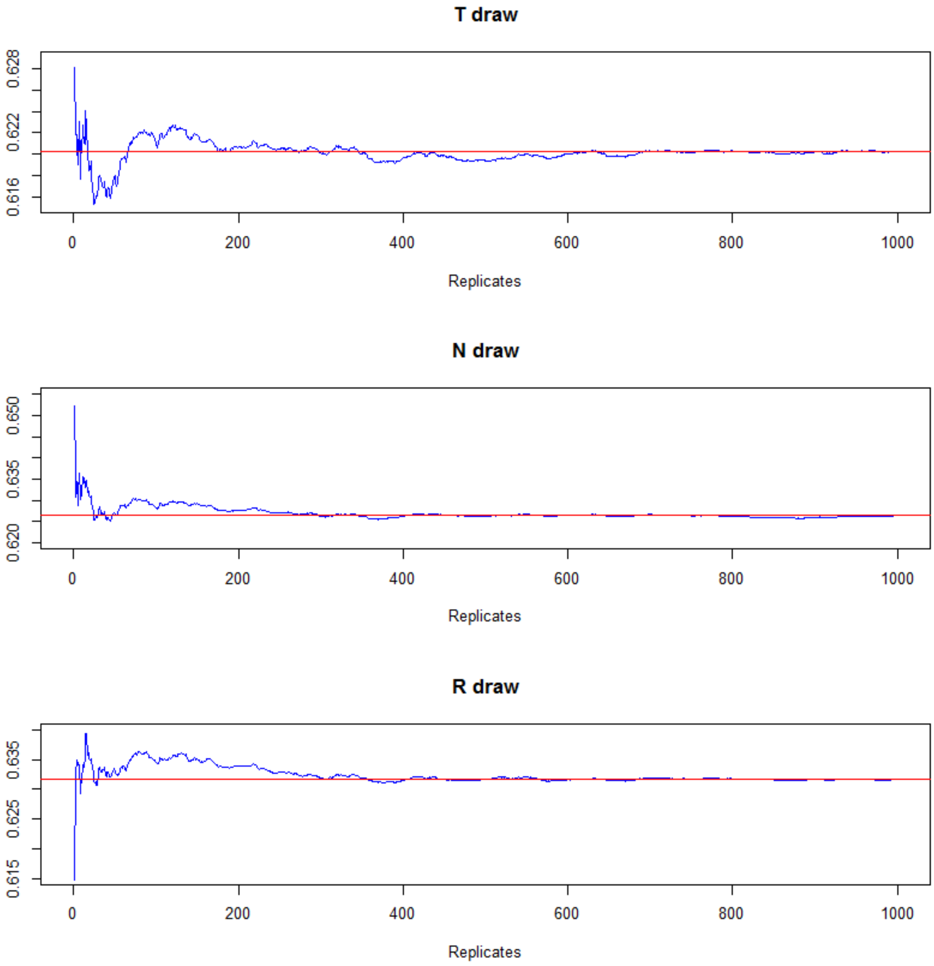

Figure 3: Density probabilities for each team under the system draw T.

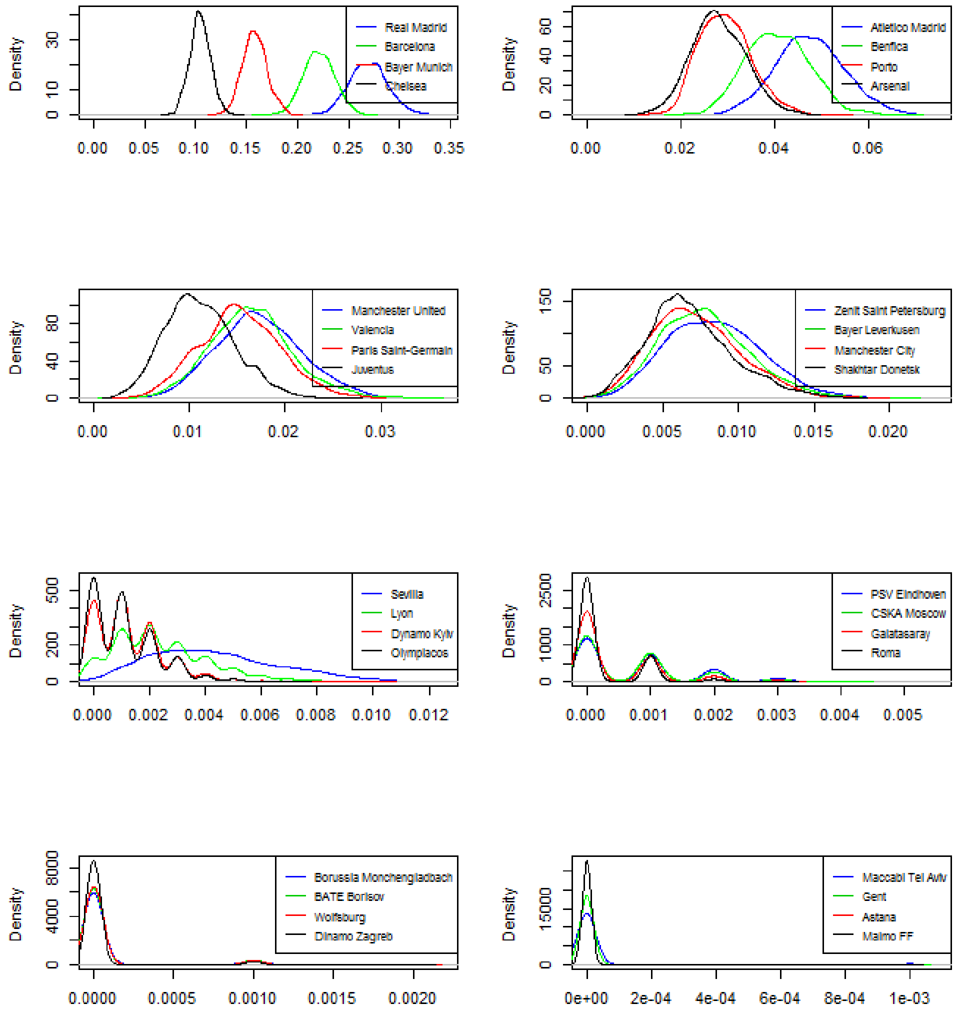

Figure 4: Density probabilities for each team under the system draw N.

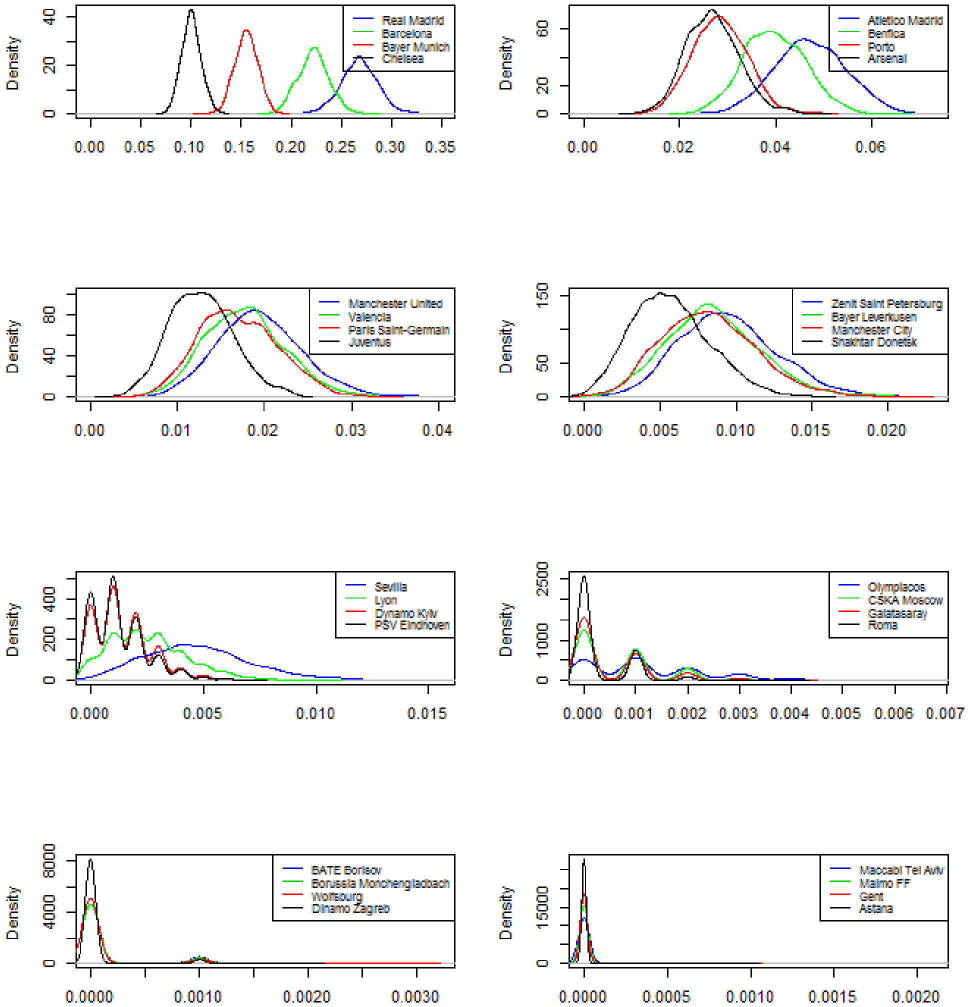

Figure 5: Density probabilities for each team under the system draw R.

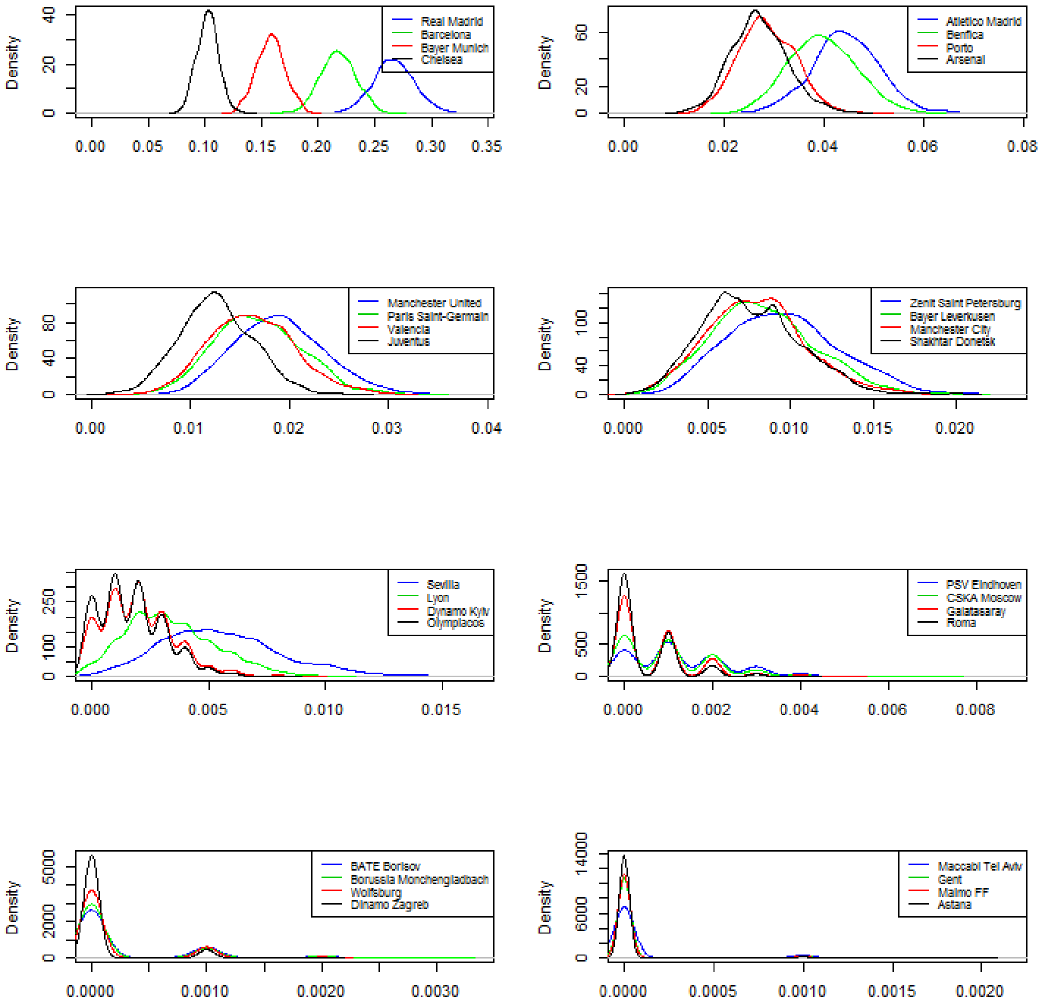

Figure 6: Density of the entropies under each system draw. Blue color is T, green color is N and red color is R.

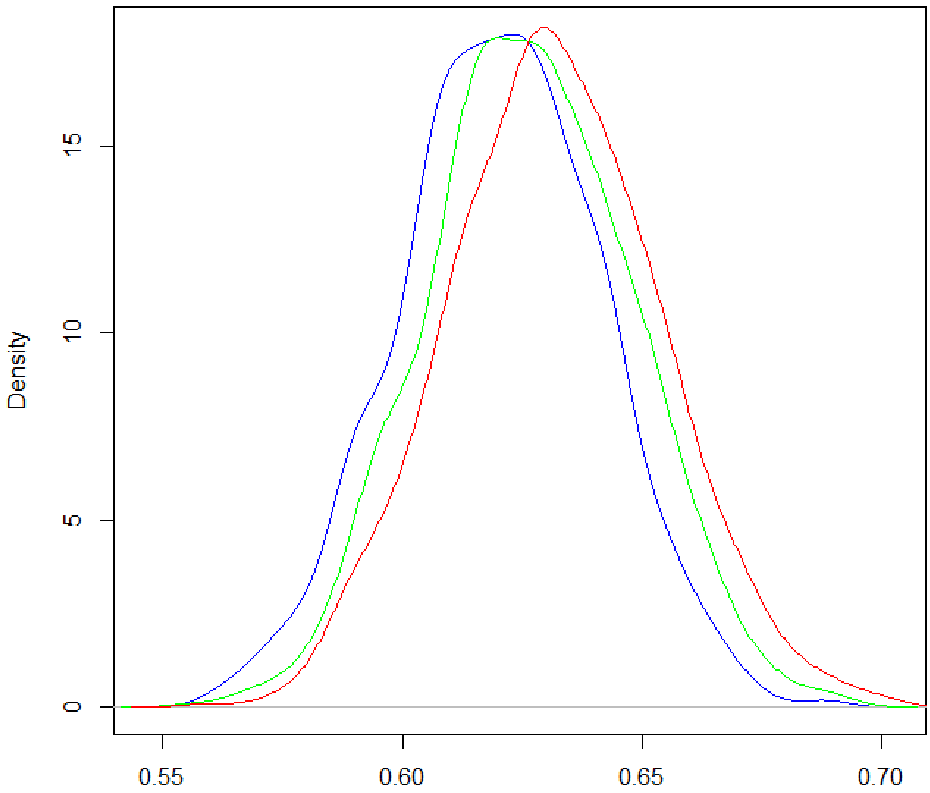

Figure 7: Plot of confidence intervals (95%) for each pair of differences of system draw on the differences of entropies between the means using Tukey's Honest Significant Difference method.

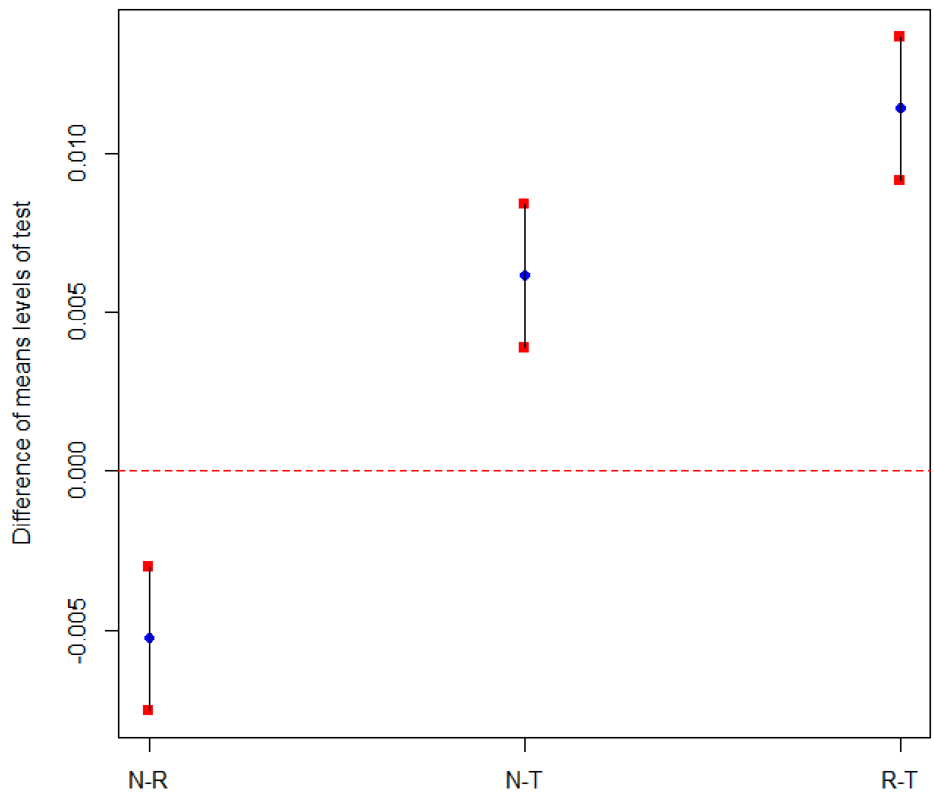

Figure 8: Plot of confidence intervals (95%) for each pair of differences T-N on the differences of percentage of qualification to round of 16 between the means of each team using Tukey's Honest Significant Difference method.

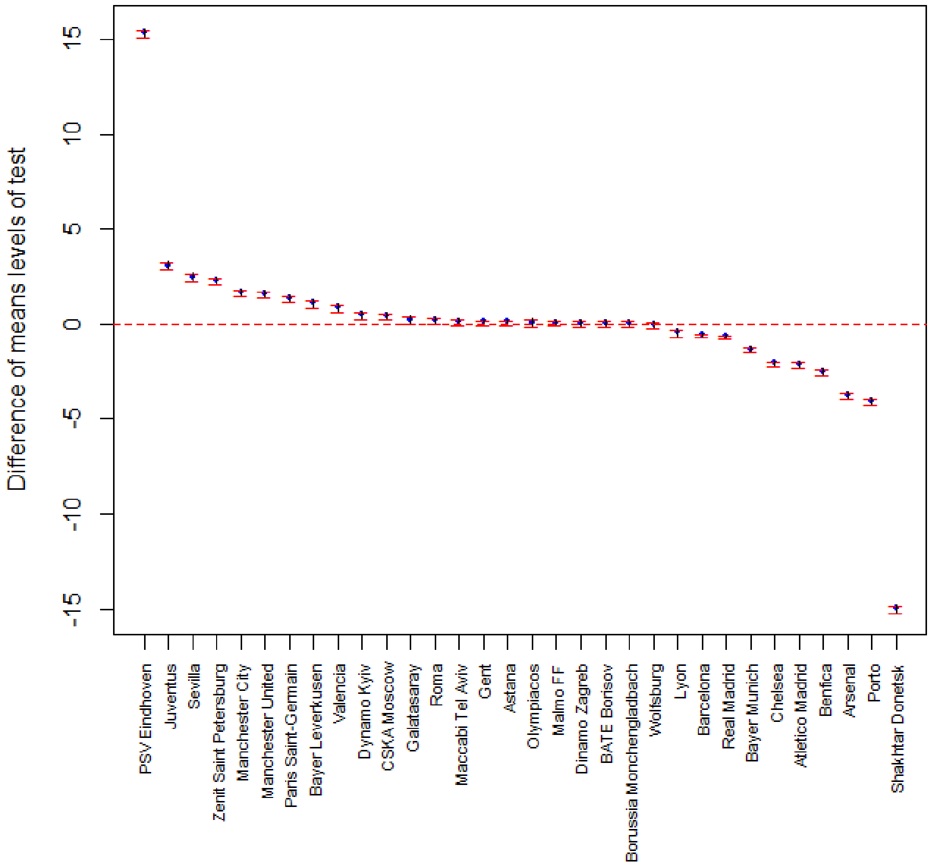

Figure 9: Plot of confidence intervals (95%) for each pair of differences T-N on the differences of percentage of qualification to quarter-finals between the means of each team using Tukey's Honest Significant Difference method.

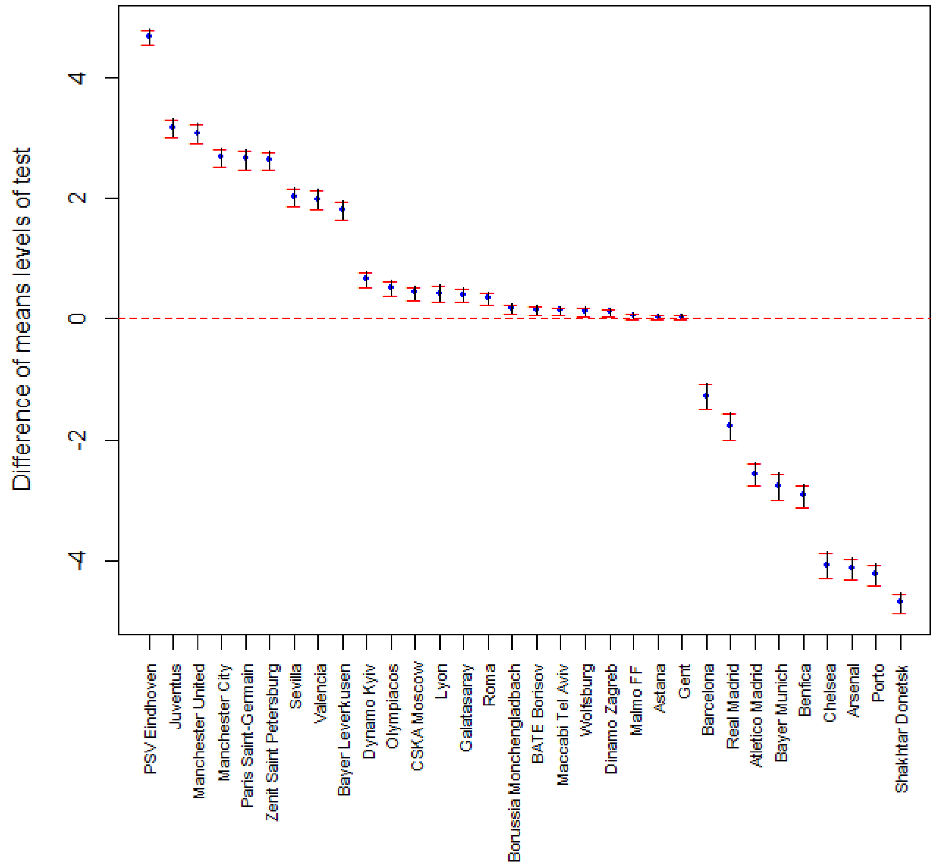

Figure 10: Plot of confidence intervals (95%) for each pair of differences T-N on the differences of percentage of qualification to semi-finals between the means of each team using Tukey's Honest Significant Difference method.